A Literature Review of Underwater Image Detection

Yawen Wang , Qinxiao Wang, Shasha Jin, Wei Long[*1], Lingxi Hu
Zhejiang Province Key Laboratory of Smart Management & Application of Modern
Agricultural Resources, School of Information Engineering, Huzhou University,
Huzhou, Zhejiang, China

Abstract. As one of the world's few large-scale aquaculture countries, the mechanization, output and efficiency are low in Chinese aquaculture industry.with the development of industrial automation technology and computer image processing technology, aquaculture will be automated, scientific management transformation. As a specific application of computer recognition technology, underwater image recognition technology is important to promote aquaculture industry automation and intelligence. In this paper, first of all introduces the existing research of underwater target image recognition, mainly presents the underwater image recognition technology based on deep learning. And then the current problems of underwater image recognition is summarized. Besides, the future development of underwater image recognition technology and its application in wisdom fishery is also discussed, which provides other scholar with some theoretical help and a comprehensive reference.

Keywords. Underwater image recognition, Deep learning, Target recognition

1. Introduction

Since the end of 1950s, with the deepening reform of the Chinese economic system and adjustment of the industrial structure, the aquaculture industry has developed rapidly [1-3]. However, Chinese aquaculture is not proportional to the status of a large country given its low degree of automation. In recent years, computer image processing technology has become more and more mature, with concepts such as smart fishing and precision agriculture. To promote the development of underwater image recognition technology and The application of underwater image recognition technology in aquaculture will greatly promote the resolution of two problems which limit the development of technology resources and promote the development of industrial automation. Due to the slow start of the research and application of underwater image recognition technology in production, the underwater image acquisition and imaging process will be affected by a variety of factors complex, causing severe noise pollution, and unclear image background photo, which is difficult for underwater image recognition processing. Based on the analysis of existing research on underwater image target recognition, this article summarizes traditional algorithms and deep learning algorithms for underwater image target recognition. Underwater image recognition problems were summarized and prospects for the development and application of smart fishing were

---

[1] *Corresponding Author:lw@zjhu.edu.cn

explored to provide references for the implementation of precision culture, structural adjustment industrial and modernization.

2. Underwater image detection

At present, many international universities, national institutes and domestic research institutes are carrying out research in this area. Domestic scholars are also committed to this research. Duan et al. [4] analyzed the research progress of fish size, shape, color and other aspects from the perspective of computer vision, focusing on fish, and summarized specific improvement measures such as image acquisition, contour extraction, feature calibration and calculation. The application of computer vision measurement in diagnosis, recognition and classification of aquatic animal diseases was analyzed and discussed. Peng et al. [5] analyzed the advantages and disadvantages of traditional and deep learning underwater image preprocessing methods, compared and summarized related methods, and analyzed the improvement of deep learning methods and the existing problems in practical application.Yu et al. [6] studied the underwater imaging characteristics affected by light, used several different image processing algorithms to extract invariant features from underwater images, and verified the feasibility of underwater target recognition through underwater red ball experiment. Yu et al. [7]paired fish, shrimp, sea cucumber and other aquatic animal data collection technology, underwater target detection data set comparison, underwater image data preprocessing method, comparison of different underwater target recognition research situation, deep learning application in recognition and The tracking aspect is sorted out.

Underwater image recognition research progress are mainly traditional manual features and deep learning during the period of time, the early scholars of underwater image recognition technology research mainly traditional manual characteristics, and is applied to the actual aquaculture environment of underwater image recognition technology is mainly the technology, but the recognition algorithms rely on traditional manual design of feature extraction, It requires relevant professional knowledge and complex algorithm debugging process, which is not universal and has low identification accuracy. With the emergence of R-CNN, deep learning began to develop rapidly in object recognition and detection, and a group of scholars emerged to apply deep learning to underwater image recognition research, and the technology gradually matured. At present, the research of underwater image recognition mainly focuses on deep learning and improving the universality and accuracy of traditional algorithms.

2.1. Period of traditional manual features

The underwater image recognition process in the traditional manual feature period is divided into regional selection, feature extraction and classifier classification, as shown in Figure 1.
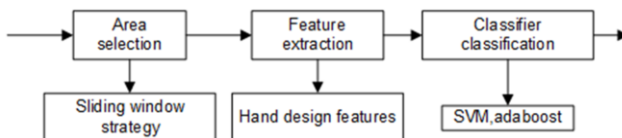

Figure 1. The process of target recognition in traditional manual feature period

Underwater image recognition is a practical application of target recognition [8], which also includes feature extraction of images. According to the actual situation under water, features such as color, texture, shape, etc. are designed and extracted, and for a special target (or one Two types of targets) are distinguished from other targets (or other types of targets), and the structure of the entire image is analyzed. Common feature extraction methods are shown in Table 1.

1) Extraction of color features

Color feature describes the attributes of all pixels in an image or an image region according to the features of pixels. Color channel histogram is the most commonly used method of color feature table. To solve the impact of natural light and artificial light on underwater imaging quality, Guzin et al. [9] adopted the underwater image enhancement technology of adaptive histogram equalization and layered difference representation with limited contrast, combined the global and local contrast enhancement technology, and carried out local color correction for underwater images. In the module, the contrast-enhanced image is converted to HSV space, and the color values of the S and V components of the image are improved by CLAHE, which solves the problem of relatively low underwater image quality. Gu et al. [10] made use of the advantages of HSV color space to make the boundary effect between background peak and data peak of H-component histogram obvious, and made the detection rate of underwater targets of the proposed EGTTSA algorithm up to 96% under different illumination conditions.

2) Texture feature extraction

Texture feature is a statistical calculation of the region containing multiple pixels, which mainly adopts image gray difference histogram and gray co-occurrence matrix for feature extraction. In order to solve the classification problem of 14 categories of freshwater fish in actual production, Liang et al. [11] extracted the fish body image from the cut fish body image into the fish body texture feature through pre-processing operation, and combined the three texture extraction methods of GLGCM, LBP, and Gabor with those of GLGCM, LBP and Gabor. The shape is combined to train 7 traditional machine learning classifiers, and the recognition rate of GLGCM texture and shape input to the random forest classifier reaches 99.52%.Yuan et al. [12] obtained texture features of fish by using grayscale co-occurrence matrix, calculated movement information of fish by optical flow method, and detected abnormal(abnormal swimming) behavior of fish by LightGBN model under automatic water quality with an accuracy of 98.5%.

3) Geometric feature extraction

Geometric features refer to the edge, position, corner and area of the target in the image, which mainly adopts feature extraction methods such as shape, contour, region and position. Jiuyi Le et al. [13] only pay attention to the contour information in the image, and use two adjustable parameters as the index of the absolute difference within the class and the class probability. In this paper, an improved OTSU algorithm is proposed, which combines the class probability and histogram gray interval, and provides a new idea for solving the fish partial overlap problem. Yan et al. [14] proposed a kernel-based K-nearest neighbor algorithm to map the data set into a high-dimensional space by optimizing the kernel function and partition it to solve the problem that the underwater image is affected by the special underwater environment, which results in multiple feature types and nonlinear characteristics. The experimental results show that the recognition accuracy of K-NN classifier based on kernel function optimization is greatly improved compared with the classifier before optimization, and the speed of K-NN classifier is more than twice that of BP neural network classifier.

Due to the complexity and diversity of the underwater environment, some scholars often use multiple feature extraction methods. In the process of underwater image recognition, single feature extraction has the advantages of small calculation amount and fast speed, but in complex situations, the ability to extract invariant features is weak. The recognition rate of adopting multi-feature extraction in subsequent experiments is relatively improved, but it has problems such as large amount of calculation and high complexity, and its real-time effect in practice is not ideal. Traditional algorithms have certain research results in solving the problem of uneven illumination and feature extraction in the underwater image imaging process. However, the recognition effect is low when the underwater image is blurred due to the artifacts caused by suspended solids and biological movement in the water. The recognition process of traditional algorithms mainly relies on artificially designed extractors, which have weak universality.

Table 1. Common feature extraction methods

Project	Method
Colour Feature	Image color channel histogram features
Textural Feature	Feature extraction based on histogram of gray difference of image、Feature extraction based on image grayscale co-occurrence matrix
Geometric Feature	Shape contour feature、Shape region feature and Position feature

2.2. Deep learning period

At present, deep learning algorithms are mainly used in underwater image extraction technology, and the research trend is on the rise. Representative algorithms include CNN, AlexNet, VGG16, ResNet, Faster R-CNN, YOLO and other models. However, due to the specificity of underwater environment, It is often necessary to improve, combine and optimize the model to obtain better results. Deep learning application of underwater image target recognition [15] general process as below in figure 2.
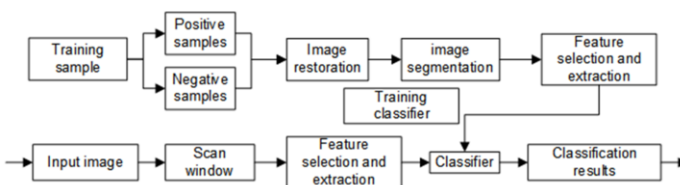

Figure 2. Target recognition process in deep learning feature period

The underwater image recognition model training samples including positive samples and negative samples, which are case sample is waiting for the underwater target sample, the negative samples refers to the other does not contain any images of the target, all the sample pictures after pretreatment and segmentation, normalized to the same size for feature selection and extraction, And get the result by classifier.

Main algorithms of deep learning applied to underwater image recognition:

1）CNN

CNN uses the deep convolutional network trained on the data set to extract image features. Since then, deep learning has developed rapidly in object recognition. AlexNet

network has 60 million parameters and 65,000 neurons, with five layers of convolution and three layers of fully connected network. The final output layer is the1000 channels of SoftMax. AlexNet uses two GPU for calculation, which greatly improves the computing efficiency.Figure 3 is the CNN network architecture diagram.Song et al. [16] used Mask R-CNN and transfer learning to identify underwater sea urchin, sea cucumber and starfish images enhanced by MSRCR algorithm to solve the problem of small number of underwater image samples and poor quality. Compared with SSD and YOLOv3 models, the recognition effect was significantly improved, providing a reference for the scientific management of aquatic product aquaculture practice. In order to solve the problem of real-time processing of underwater acquired images, Chen et al. [17] integrated the field programmable gate array into the CNN model and realized the division of the safe area and non-area of the underwater environment. The experimental processing results are consistent with the PC end, which is of great significance for underwater environment exploration. Jin et al. [18] such as underwater image recognition speed, in the feature extraction of CNN model, integrated into the pyramid mechanism and significance test method, the improved model is able to automatically extracting effective features from a large number of underwater image data set, and improved model, compared to its at the same time, improve the recognition rate prediction speed. At present, CNN model combines underwater image processing with multi-scale attention mechanism, transfer learning, BOF[19], pyramid mechanism, etc. In order to solve the problem of poor recognition effect of traditional machine learning algorithm in complex environment, Xue et al. [20] optimized AlexNet model by reducing part of redundant convolution layer and combining the flexible attentional algorithm based on terms and transfer learning. The average recognition rate of four fish species, including crothian fish, sinkfish, blue lip fish and dovetail fish, reached 97%. Ju et al. [21] integrated the flexible attentional mechanism based on terms into AlexNet model, and then migrated the pre-training weight of ImageNet data set to the optimized network. Compared with the existing research methods, the training time is short and the accuracy is high. In order to realize the automatic classification of fish species, Yuan et al. [22] input the pre-processed data of five fish species into the fine-tuned AlexNet network for training and extract the high-semantic features. In addition, they input the high- semantic features into the classifier for training through transfer learning, which not only improves the accuracy but also greatly shortens the training time. Abinaya N.S.[23] et al., in order to solve the changes in body structure and direction caused by fish out of water, input the images of fish head, fish scale and fish body respectively into AlexNet network for training using transfer learning, and introduce naive Bayesian fusion deep learning network layer. In the FishPak data set, Compared with the non-converged network layer, the classification accuracy is increased to 98.6%. In a single fish and simple video background, it has a better effect by integrating attention mechanism and transfer learning algorithm.

Its optimization model has been applied to underwater image recognition research covering many kinds of aquatic creatures, underwater environment images, sonar images, etc. The research technology of recognition efficiency and accuracy has gradually matured.
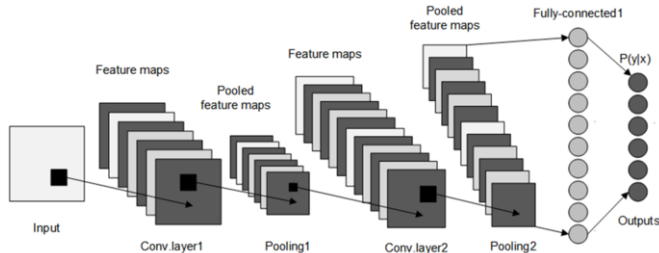
Figure 3. CNN

2）VGG16

VGG16 simplifies the neural network structure and adopts 3x3 convolutional kernel.The network architecture is shown in Figure 4 below. In order to solve the problem of image blurring captured from the video, Dong et al. [24] integrated spatial attention mechanism and layered simplified bilinear features into VGG16, and obtained better recognition rate in F4K data set after extracting high semantic features. In order to solve the complexity of traditional algorithm extraction features, Yang et al.[25] added a level of regularization, optimized activation function and migration weight to optimize the VGG16 model, and realized the accurate recognition and classification of silver carp, black fish, yellow catfish, carp and tilapia. In order to realize dense fish swarm detection, Wang et al. [26] removed the full connection layer of VGG16 and combined it with the voids convolutional neural network to expand the perception domain and give attention to the resolution at the same time. It has a good identification effect for determining whether the fish swarm density is reasonable in the fixed point area. In fish swarm classification and complex video background, it is better to integrate attention mechanism and transfer learning algorithm.
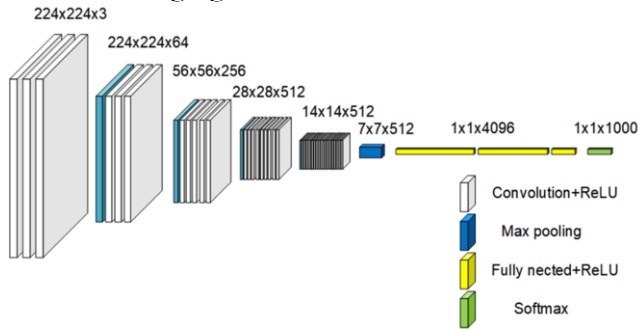

Figure 4. VGG16

3）Faster R-CNN

The Faster R-CNN network designs a candidate region generation network and implements end-to-end and quasi-real-time deep learning detection by using multiple reference Windows. The network architecture is shown in Figure 5 below. In order to solve the problem of insufficient underwater fish data, Yuan et al. [27] proposed to carry out secondary transfer learning based on Faster R-CNN, make use of high-definition fish data set for the first migration, and use the fine adjustment network of underwater fish images to improve the recognition effect of small-scale underwater fish data sets. In order to solve the problem of background diversity of Marine targets, Wen et al. [28] integrated

the attention mechanism into the Faster R-CNN network and combined the convolutional network with SKNet network, and achieved good results in the recognition of complex Marine targets.
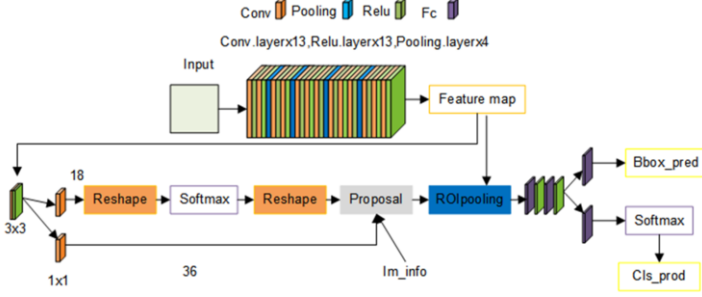

Figure 5. Faster R-CNN

4）YOLO

YOLO takes the entire image, propagates it forward at once, and uses it to study underwater images to speed up recognition. The network architecture is shown in Figure 6 below. In order to solve the problem of accuracy and efficiency of small target recognition in the unknown underwater environment, Yuan et al. [29] used images of sea cucumber, sea urchin and scallop collected underwater to train the fine-tuning YOLOv3 algorithm, and the target recognition rate reached 99%. In order to eliminate the influence of underwater light refraction on the measurement of the size of sea cucumbers, Dong et al. [30] used visual correction and Gaussian model to build the segmentation algorithm of sea cucumbers image to find the optimal point for the measurement of the size of underwater sea cucumbers image and better realize the measurement of the size of sea cucumbers in the underwater environment. Hu et al. [31]in order to solve the waste of feed particles in aquaculture and reduce the problem of high ammonia nitrogen content caused by uneaten feed dissolved in water and affecting the quality of fish, modified the connection mode of feature FPN network and PANet, and replaced the feature mapping in the original YOLO-V4 network responsible for a wide range of features with a more fine-grained feature mapping. By removing redundant operations, the identification accuracy of YOLO V4 in actual breeding environment is improved to 92.61% compared with the original YOLO V4, which provides a new idea for scientific feeding.
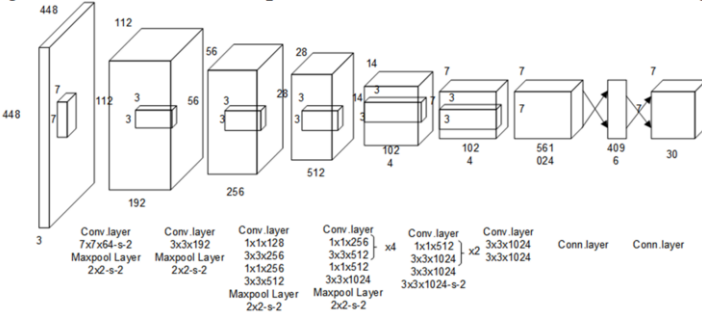

Figure 6. YOLO

The development of underwater image recognition algorithm is summarized from the aspects of development stage method, significance and thought, as shown in Table 2.

Tab.2 Development of underwater image recognition algorithm

Stages of development	Method	Significance	Thought
Period of traditional manual features	Manual feature selection, Machine Learning	Emancipate labor force and provide new ideas for industrial development	Sliding window detection, feature selection algorithm, multi-stage processing strategy, multi-scale pyramid
Deep learning period	CNN	Since then, the application of deep learning in recognition has developed rapidly	Train the network on the data set to extract image features
	AlexNet	Improvements were made in LeNet to improve efficiency	Convolution followed by full join
	VGG16	Improvements to AlexNet	Several convolution kernels of 3x3 are used to replace its large convolution kernels
	Faster R-CNN	R-CNN and SPPNet are improved. Fast R-CNN is improved to realize end-to-end deep learning detection	Candidate region generation network
	YOLO	Compared with previous algorithms, it is faster to be applied to projects in real time	The whole picture is input and propagated forward at once

3. Problems and challenges

Though years of underwater image target recognition extensively studied and developed a large number of techniques and algorithms, the robustness of recognition methods, accuracy, efficiency and scope of all had the very big enhancement, but the traditional algorithm and deep learning algorithm is applied to actual production there are still some shortage, mainly reflected in the following respects.

1) in the process of experiment data acquisition, will be affected by uneven illumination, the turbidity of water environment, the underwater camera protection block, shooting Angle different, detecting target size, speed and other factors, will cause the collected experimental data can't provide enough information reduction target, target detection result is not stable, test results do not tally with the expected.

2) There are few experimental data in the real water environment, and the cost of underwater photography is high. It is difficult to collect experimental data set required for deep learning pre-training, and the training model cycle is long, resulting in poor real-time identification effect.

3) At present, humans do not know how to clearly identify high-dimensional objects from low-dimensional, which cannot provide direct guidance for the research of underwater image object recognition. In addition, the establishment of underwater image knowledge base requires not only the knowledge of the color, texture and shape of the object itself, but also the knowledge of the relationship between objects. The preliminary preparation of knowledge base will directly affect the accuracy of target recognition in the later stage of the project.

4) At present, most of the underwater image processing visual systems built by people are only applicable to specific environments or specific objects, and the recognition model is not universal.

4. Prospect

The recognition and detection of underwater images are the prerequisites for the application of computer technology to aquaculture playing an important role in the development of fishing, precision and the promotion of structural adjustment of aquaculture In the future, underwater image recognition technology could be explored in the following areas:

1) Quantity of high quality underwater images and experimental data. High quality images are an important part of recognizing underwater images. High quality images can improve image preprocessing speed and identification accuracy. To study efficient underwater image data acquisition technology, develop and improve underwater image quality technology, and solve the problem of low amount of high-quality experimental underwater image data. To promote the development of a dataset sharing platform for underwater image recognition and the advancement of underwater image recognition technology.

2) Key technological breakthroughs. A combination of various technologies is adopted to resolve the contradiction between real-time and the precision of the underwater image recognition algorithm to improve the efficiency of the algorithm applied in real-life aquaculture.

With the development of underwater image recognition technology, more and more new technologies will be applied to smart fishing, which will have a profound impact on the whole fishing culture system and the lives of people.

5. Acknowledgement

References

[1] YU Ning et al,Study on the development status and countermeasures of intelligent fishery,Chinese Fisheries Economics 39 (2021),13-21.

[2] XIA Yingkai et al,Progress of underwater robots for aquaculture,Journal of Huazhong Agricultural University(2021),1-13.

[3] Wang Ke et al,Research on underwater flexible target recognition algorithm under non-uniform light. Intelligent Service Robotics(2020),1-10.

[4] Duan Yan'e et al,Review on visual characteristic measurement research of aquatic animals based on computer vision,Transactions of the Chinese Society of Agricultural Engineering(2015),1-11.

[5] PENG Xiaohong et al,Review of Underwater Image Preprocessing Based on Deep Learning,Computer Engineering and Applications(2021),1-14.

[6] WU Yu et al,Research on the Underwater Optical Imaging Processing and Identification,Ship Electronic Engineering(2019),93–96.

[7] YU Hong,Research progress on object detection and tracking techniques utilization in aquaculture:a review,Journal of Dalian Ocean University(2020) ,793–804.

[8] LIU.Fet al. Underwater Target Recognition Using Convolutional Recurrent Neural Networks with 3-D Mel-Spectrogram and Data Augmentation. Applied Acoustics, (2021), 178: 107989.

[9] ULUTAS G et al,Underwater Image Enhancement Using Contrast Limited Adaptive Histogram Equalization and Layered Difference Representation. Multimedia Tools and Applications(2021).

[10] GU Jialu et al,Research on Threshold Segmentation Algorithm of Underwater Robotic Fish Based on Edge Gravity Center Template,Machinery & Electronics(2018),25-28+32.

[11] LIANG Zhaodong et al, An automatic method for freshwater fish species classification using shape and texture features,Journal of Nanjing Agricultural University(2021),576-585.

[12] YUAN Hongchun et al,Detection of fish abnormal behavior based on LightGBM model,Fishery Modernization(2021),47–55.

[13] Jiuyi Le and Lihong Xu. An Automated Fish Counting Algorithm in Aquaculture Based on Image Processing[C]. Shenzhen, China, 2017.

[14] YAN Liangtao et al,Application of k-NN based on Kernel in underwater target recognition,Journal of Applied Acoustics(2019),448-451.

[15] WANG Sheng-gui et al,Research on underwater target recognition method based on deep learning,Ship Science and Technology(2020),141–145.

[16] Song Shaojian et al,Object recognition research of underwater creature based on Mask R-CNN and transfer learning ,Application Research of Computers (2020)(S2 vo 37): 386-388+391.

[17] CHEN Kaifeng et al,Underwater target recognition system based on FPGA and CNN,Transducer and Microsystem Technologies(2021), 103-105+109.

[18] JIN Leilei et al,Sonar image recognition of underwater target based on convolutional neural network,Journal of Northwestern Polytechnical University(2021),285-291.

[19] QUAN Wenwen et al,Algorithm of underwater target recognition based on CNN features with BOF,Journal of Shandong University(Engineering Science)(2019), 107–113.

[20] XUE Yongjie et al,Fish Recognition Algorithm Based on Improved AlexNet, Electronic Science and Technology(2021),12–17.

[21] Ju Zhiyong and Xue Yongjie. Fish species recognition using an improved AlexNet model, Optik(2020), 223 : 165499-.

[22] Yuan Xiaoping et al,Research on Fish Recognition Method Based on Transfer Learning,Computer Technology and Development(2021), 52–56.

[23] N.S. Abinaya and D. Susan and S. Rakesh Kumar. Naive Bayesian fusion based deep learning networks for multisegmented classification of fishes in aquaculture industries. Ecological Informatics,(2021), 61(prepublish).

[24] DONG Shaojiang et al,Fish recognition based on hierarchical compact bilinear attention network,Computer Engineering and Applications(2021),1-6.

[25] YANG Chun-lan et al,Freshwater fish classification recognition based on convolutional neural network,Journal of Southwest Minzu University(Natural Science Edition)(2020), 611–618.

[26] WANG Jinfeng et al,Experimental research on fish density detection based on improved deep learning model,Fishery Modernization(2021),77-82.

[27] YUAN Hongchun et al,Detection of underwater fish based on Faster R-CNN and image enhancement,Journal of Dalian Ocean University(2020), 612–619.

[28] WEN Lili et al,Ocean target recognition algorithm based on attention mechanism and Fast R-CNN,Journal of Dalian Ocean University(2021),1-15.

[29] YUAN Li-hao et al,Small Underwater Target Recognition Based on YOLOv3.Ocean Engineering Equipment and Technology(2018),118-123.

[30] DONG Peng et al,Automatic Measurement of Underwater Sea Cucumber Size Based on Binocular Vision,Computer Engineering and Applications(2021),271-278.

[31] Hu Xuelong et al. Real-time detection of uneaten feed pellets in underwater images for aquaculture using an improved YOLO-V4 network. Computers and Electronics in Agriculture(2021) , 185.